© 1993 Oxford University Press

Nucleic Acids Research, 1993, Vol. 21, No. 5 1087–1095

Cloning of the zebrafish *krox-20* gene (*krx-20*) and its expression during hindbrain development

Elli Oxtoby and Trevor Jowett*
Department of Biochemistry and Genetics, Medical School, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4HH, UK

Received December 22, 1992; Accepted January 18, 1993 EMBL accession nos X70322, X70323

ABSTRACT

To begin to examine the function of genes that control early development in the hindbrain, we have screened an embryonic zebrafish cDNA library with a murine *krox-20* gene probe that contained the conserved zinc finger regions. We have isolated two overlapping cDNAs, zf187 and zf201 which are homologues of the murine *krox-20* gene. The N-terminal of the longest cDNA (zf201) contains two acidic regions identical to those of the murine *krox-20*. This indicates that the functional organisation of these proteins is probably conserved. Northern Blot analysis identified a single transcript of 2.0 kb. Wholemount *in situ* hybridisation established that expression of the zebrafish gene (*krx-20*) first appears at 100% epiboly as a single anterior domain of the prospective neuroepithelium, followed very soon after by a second more posterior domain. The alternating pattern of expression of this gene in rhombomeres(r) r3 and r5 is apparent by 12 hr post-fertilisation, that is prior to the morphological appearance of the rhombomeres. Around 14 hr neural crest migration begins from the dorsal surface of r5, moving caudally into r6 and then ventrally towards the pharyngeal arches. Crest migration is not apparent at or after 16 hr. No neural crest migration was observed from r3. Expression of *krx-20* is down regulated firstly in r3 around 26 hr and later in r5 around 30 hr.

INTRODUCTION

The process of segmentation is a widely employed strategy in development that, allows the generation of similar classes of cells in adjacent territories but with distinct positional identities. Recent studies of both the zebrafish (1) and the chick (2) central nervous system demonstrated that the hindbrain is organised on a segmental basis. During an early phase of development the hindbrain is divided into seven repetitive structures, called rhombomeres, with intervening rhombomere boundaries. Clonal analysis using intracellular markings indicated that from the earliest morphological appearance of rhombomeres cell mixing is restricted at their boundaries (3). Thus rhombomeres are units of cell lineage restriction in the same way as compartments in the imaginal discs of insects (4). Neurogenesis is initiated in alternating rhombomeres (r), (r2, r4 and r6) and the Vth, VIIth and IXth cranial nerves arise from successive pairs of rhombomeres r2 and r3, r4 and r5, and r6 and r7 respectively indicating the existence of mechanisms operating at two segment intervals (2).

This segmental organisation of the hindbrain raised the question of the existence of a matching pattern of gene expression which would control the process of segmentation and define the identity of each unit. The observation that the murine *krox-20* gene is expressed in two rhombomeres revealed a candidate for such a segmentation gene and provided molecular evidence for hindbrain metameric organisation (5). In addition a number of genes including homeobox genes were shown to have limits of expression which coincide with rhombomere boundaries (6), suggesting that these genes may be part of a regulatory network governing pattern formation in the hindbrain.

krox-20 encodes a protein with three C_2H_2-type zinc fingers (7). This protein was shown to bind to a specific DNA sequence and to act as a transcription factor (8). It belongs to a small subfamily of proteins, with similar zinc fingers, which recognise identical or very closely related GC-rich sequences. Three other members have been identified *krox-24 (Egr-1, Zif268, NGFI-A and TIS8* (9–13)), *EGR-3* (14), and *NGFI-C* (15). These proteins are very closely related within their putative DNA-binding domains, but they appear to be much less conserved elsewhere.

Zinc finger coding sequences have been found in a number of genes that control *Drosophila* development. Mutants for the *Drosophila* gap gene *Krüppel* lack all thoracic and several abdominal segments (16)—another *Drosophila* gap gene, *hunchback*, controls development of head structures (17). The *Drosophila* genes *Serendipity β* and *Serendipity δ*, which result in an embryonic lethal phenotype when mutant, also contain zinc fingers (18) as does *ovo*, a *Drosophila* gene that is required for the development of the female germ line (19).

Although direct evidence of the role of *krox-20* is not as yet available, its expression pattern in the hindbrain suggested a possible interaction between *krox-20* and homeobox-containing genes. Thus *krox-20*, in combination with other genes, could regulate the high level expression of *Hox-2.7* and *Hox-2.8* in r5 and r3-r5 respectively and the restriction of *Hox-2.9* expression

* To whom correspondence should be addressed

to r4 (50). This idea is consistent with the finding that *krox-20* protein binds to sequences in the 5′ flanking genomic region of the *Hox-1.4* gene *in vitro* and can utilise this site to activate transcription *in vitro* (8).

A full understanding of the function of genes that control early development of the hindbrain requires a combination of molecular and mutational analysis. The zebrafish (*Brachydanio rerio*), is believed to be a good system for studying vertebrate development because it provides the possibility of combining excellent embryology with fairly good genetics. The zebrafish is particularly attractive because females can produce a large number of eggs; the embryos develop outside the mother; in addition their early development is rapid (gastrulation takes 5 hr) and because the embryos are transparent direct observation of their embryonic development is possible; the generation time is 60−70 days, and large numbers can be raised in a relatively small space. Screening for recessive mutations is facilitated by methods for the production of haploid and gynogenetic offspring (20,21). Furthermore, it is possible to produce transgenic zebrafish via microinjection of foreign DNA into the cytoplasm of fertilised eggs (22,23). Despite these attractions however the zebrafish model lacks many basic tools. There is no linkage map, the number of cloned genes is small and so is the number of accurately described mutants.

In this study we report the cloning of *krox-20* from zebrafish (*krx-20*). We have used *in situ* hybridisation to show that its expression pattern is conserved suggesting that its role in hindbrain development is conserved.

MATERIALS AND METHODS

Library construction and screening

A cDNA library was constructed in λgt10 using poly(A)$^+$ RNA from zebrafish embryos (20−24 hr old). The cDNAs were synthesised using a cDNA cloning kit (Pharmacia) and subsequently were inserted into λgt10 arms using EcoRI-NotI adapters. 30,000 recombinant plaques were screened in duplicate with a ^{32}P labelled PstI-ApaI fragment (753 bp long) of the mouse *krox-20* cDNA (provided by Dr D.Wilkinson). The fragment contained the sequence encoding the zinc fingers and about 240 bp of the 3′ untranslated region. Plaques were lifted onto nitrocellulose filters and hybridised overnight in 5×SSC; 5×Denhardt's; 1% SDS; 50 mM sodium pyrophosphate (pH 6.8) at 60°C. The filters were washed to a final stringency of 1× SSC; 0.25% SDS, 20 min at 60°C. Positives were subjected to two rounds of plaque purification. Two cDNAs zf187 and zf201 were selected for further analysis.

PCR amplifications

Amplification of zebrafish genomic DNA, using 17-mer oligonucleotide primers derived from the cDNA sequence of zf201, revealed the presence of an intron upstream from the zinc fingers. For the sequence and position of each primer see table 1. This sequence (445 bp long) was then amplified by PCR (polymerase chain reaction) using zf87 and zf230 primers that flank the intron sequence. The PCR reaction was carried out using 500 ng of genomic DNA in a buffer containing; 50 mM KCl, 10 mM Tris (pH 9.0), 1.5 mM $MgCl_2$, 0.01% gelatine, 0.1% Triton X-100, 0.2 mM of each dNTP, 500 ng of each primer and 2.5 i.u. of Taq polymerase (BCL), in 100 μl reaction volume. This reaction mixture was subjected to the following PCR profile (using a Biometra TRI0 thermoblock): (i) 94°C, 5 min; (ii) 30 cycles of 55°C, 2 min; 72°C, 3 min; 94°C, 2 min; and (iii) a final cycle of 55°C, 2 min; 72°C, 10 min. This yielded enough DNA for cloning into pBluescript KS(+) (Stratagene). For this purpose the PCR fragments were gel purified, extracted with phenol/chloroform, ethanol precipitated and then treated with Klenow and T4 polynucleotide kinase and ligase in a single step as described by J.B. Lorens (24). Two clones from two independent PCR reactions (CG zfkrx-20-16) and (CG zfkrx-20-70) were selected for further analysis.

DNA analysis and sequencing

Preliminary restriction and Southern blot hybridisation analysis (essentially as described in Sambrook (25)) revealed that zf187 and zf201 contained inserts of 1.0 kb and 1.9 kb respectively. These were subcloned in pBluescript KS(+) (Stratagene) as, either a NotI (zf187) or EcoRI (zf201) fragments. The intron DNA was subcloned as blunt end fragments. The clones were subsequently sequenced as double stranded template in both directions using Sequenase v2.0 Kit (USB) and synthetic oligonucleotide primers, according to the manufacturer's protocol. CG zfkrx20-16 and CG zfkrx20-70 produced identical sequences and so did the two cDNA clones except that zf187 was about 900 bp shorter than zf201 at the 5′ end.

RNA isolation and Northern blot hybridisation analysis

Total RNA was extracted by the guanidinium isothiocyanate-CsTFA (caesium trifluoroacetate, Pharmacia) method essentially as described in Sambrook (25). Poly(A)$^+$ RNA was selected using biotinylated oligo(dT) and streptavidin paramagnetic particles as described in the manufacturer's protocol (PolyATtract™, Promega) mRNA isolation system. 2 μg of poly(A)$^+$ RNA was fractionated by electrophoresis on a 1.2% agarose-formaldehyde gel, and blotted onto nitrocellulose membranes. The hybridisation conditions were the same as for the library screening (above). Filters were washed at a final stringency of 0.1×SSC; 0.1% SDS at 65°C for 20 min.

Whole mount *in situ* hybridisation

The localisation of *krox-20* transcripts in zebrafish embryos was analysed by using whole mount *in situ* hybridisation. Hand dechorionated zebrafish embryos were fixed in 4% paraformaldehyde in 1×PBS at 4°C for a minimum of 2 hr and stored in methanol at −20°C. Before hybridisation stored embryos (50−100 embryos/tube) were rehydrated in a graded methanol/PBS series for 5 min at each step; followed by four washes in 100% PBTw (1×PBS, 0.1% Tween 20); 20 min in 10 μg/ml proteinase K in PBTw; rinsed 2× in 2 mg/ml glycine/PBTw; 20 min in 4% paraformaldehyde in PBS; washed

Table 1. Sequence of each primer used in genomic DNA amplifications.

Primer	Sequence	Position
zf 87 (*)	5′ACCCTCTTGCCGATAGC3′	88−104
zf 218	5′ATGAGCACGGAGAAGCG3′	219−235
zf 382	5′GATCCTGGGCATGACCC3′	383−399
zf 571	5′TCCACCTCCACCTGTCC3′	570−586
zf 230 (*)	5′CTCCGTGCTCATATCCC3′	230−214
zf 469	5′GCAGCTGAGAGTGCTGG3′	470−453
zf 705	5′ATGTCCCTCTGGCACGG3′	700−684

The nucleotide positions shown correspond to the cDNA sequence (Fig.2A) used to synthesize the corresponding primers. The two primers that flank the intron sequence (Fig.2B) are marked with (*).

5× in PBTw 5 min each. After this, the embryos were prehybridised in hybridisation buffer (50% formamide; 5×SSC; 0.1% Tween-20; 50 μg/ml heparin; 500 μg/ml tRNA; pH 6.0 with citric acid), for 1 hr at 65°C. Afterwards the hybridisation buffer was replaced by fresh hybridisation buffer containing digoxigenin labelled RNA probe at 0.5 ng/μl final concentration and incubated overnight at 65°C.

An equal number of embryos processed in parallel without hybridising to an RNA probe were used to preadsorb the anti-digoxigenin antibody. The antisera was diluted 1:400 in PBTw; 2 mg/ml BSA; 2% sheep serum. Preadsorption was carried out for 2 hr at room temperature, followed by overnight incubation at 4°C, with gentle shaking. After hybridisation the embryos were washed at 65°C shaking for 10 min each wash with: 75% hybridisation buffer, 25% 2×SSC; 50% hybridisation buffer, 50% 2×SSC; 25% hybridisation buffer, 75% 2×SSC. Then twice for 30 min each with 0.2×SSC. Followed by 5 min washes at room temperature with: 75% 0.2×SSC, 25% PBTw; 50% 0.2×SSC, 75% PBTw, and finally 100% PBTw. The embryos were incubated at room temperature shaking for 1 hr in 100% PBTw, 2 mg/ml BSA, 2% sheep serum and afterwards for 2 hr with 100μl/100 embryos of preadsorbed antibody (diluted 1:1000). The embryos are subsequently washed in several changes of PBTw lasting about 2 hr followed by 3 washes 5 min each in reaction buffer (100 mM Tris−HCl pH 9.5; 50 mM $MgCl_2$, 100 mM NaCl, 0.1% Tween-20) and stained by adding 4.5 μl/ml NBT and 3.5 μl/ml X-phosphate (BCL) until an optimal signal was obtained (about 12−20 min). The staining reaction was stopped by washing in several changes of PBTw. The embryos were kept at 4°C and photographed in 1% methylcellulose, 1×PBS, 0.1% sodium azide.

RNA probes

Antisense RNA probes were transcribed *in vitro* from PstI linearised plasmid using digoxigenin labelled UTP and T_3 RNA polymerase as described in the manufacturer's manual (DIG RNA Labelling Kit, BCL). The RNA probe was denatured at 80°C for 2 min before use. Finally 3 μl of the riboprobe was run on a Northern gel and transferred onto nitrocellulose as described above, and its labelling efficiency was detected by enzyme-linked immunoassay according to the manufacturer's instructions (DIG Nucleic Acid Detection Kit, BCL).

RESULTS

Characterisation of zebrafish *krx-20* cDNA and intron sequences

We screened a cDNA library prepared from 20−24 hr zebrafish embryos with a murine *krox-20* probe and isolated two zebrafish clones (zf187 and zf201). Both clones were sequenced and found

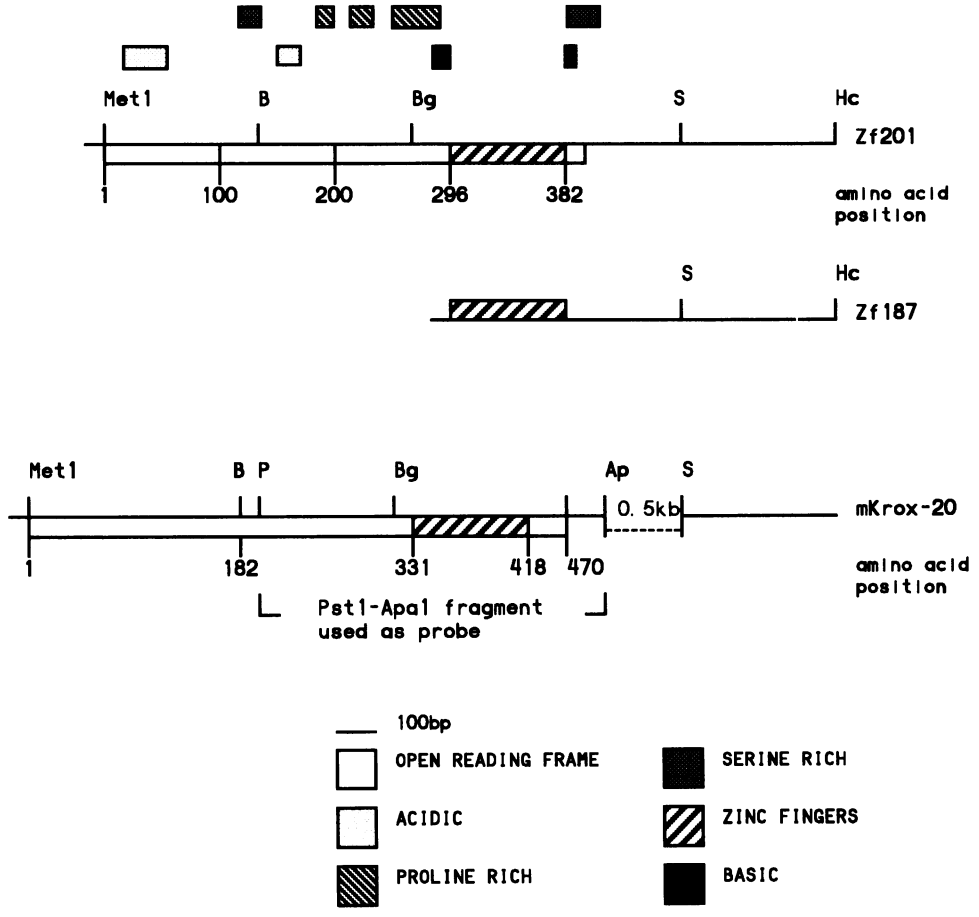


Figure 1. Schematic representation of two cDNA clones of *krx-20*. The open reading frame is shown on the longest cDNA and several features of the krx-20 protein are indicated: zinc fingers, proline-rich, serine-rich, acidic and basic regions. The mouse *krox-20* cDNA (30) is also shown and the PstI-ApaI fragment used as probe in this study is bracketed. Ap = ApaI; B = BamHI; Bg = BglI; P = PstI; S = SphI; Hc = Hinc II.

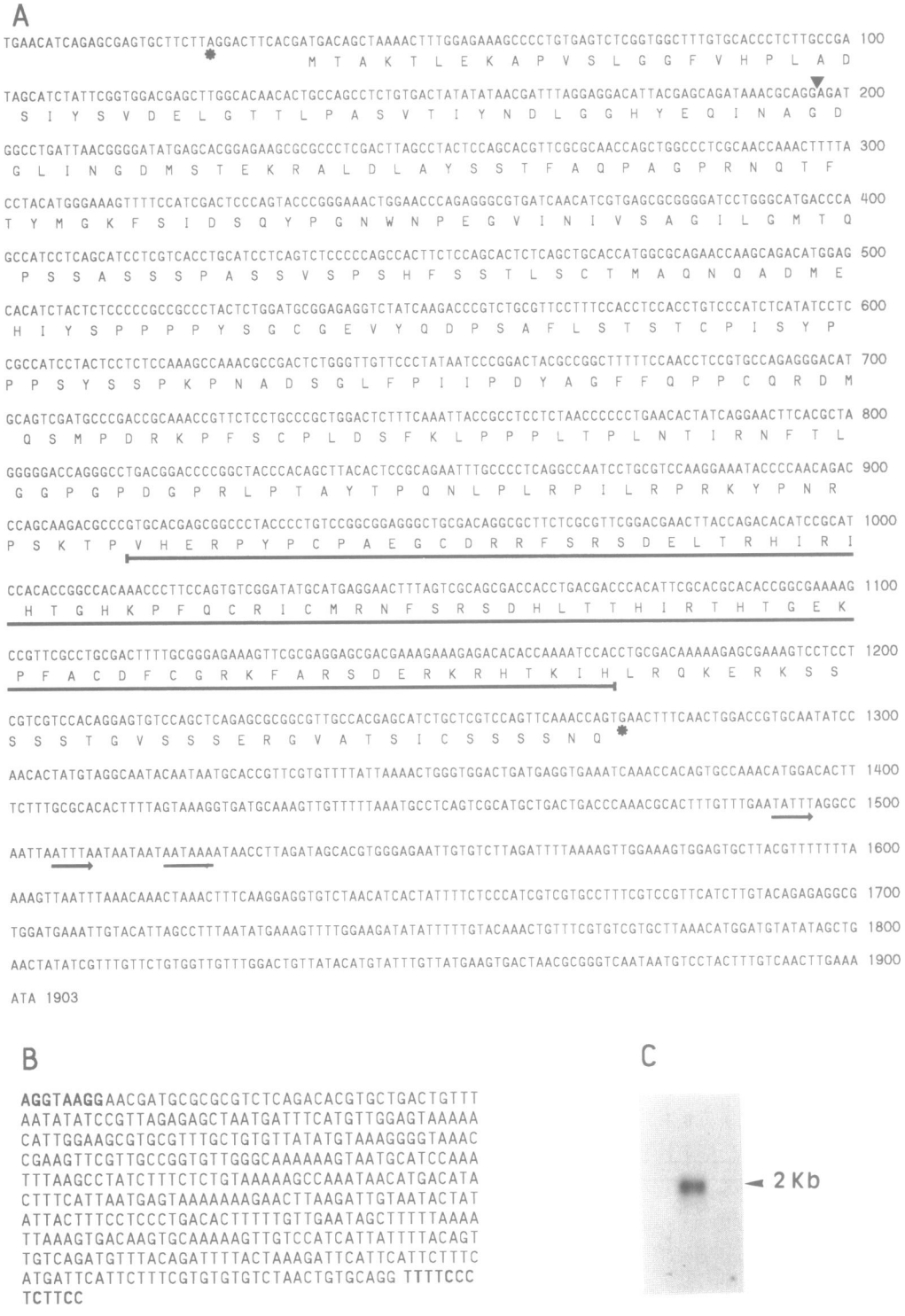


Figure 2. A: Nucleotide sequence of *krx-20* and predicted amino acid sequence of the protein. The position of the intron (sequence shown in B) is marked by a triangle. Asterisks indicate a stop codon. The consensus signal for poly(A) addition is underlined. The sequences ATTTA potentially involved in mRNA selective degradation are shown by arrows. The zinc finger area is bracketed. **B**: Intron sequence. The complete nucleotide sequence of the intron, derived as described in the text, is shown. The 5′ and 3′ splice consensus is shown in bold. **C**: Northern Blot analysis of poly(A)$^+$ RNA from 20−24 hr zebrafish embryos. A message of approximately 2.0 kb was detected using zf201 as probe. The same result was obtained when using zf187 as probe.

to be identical within their overlapping regions (fig. 1). The complete nucleotide sequence and deduced amino acid sequence of zf201 (shown to be 900 nucleotides longer at the 5′ end than zf187) are shown in fig. 2A. Northern blot analysis identified a single transcript of about 2.0 kb (fig. 2C). This suggests that zf201, though nearly a full length cDNA, is truncated at the 3′ end. Thus, the polyA tail is missing. As shown in fig. 2A the nucleotide sequence (1903 nucleotides in total) includes, 5′

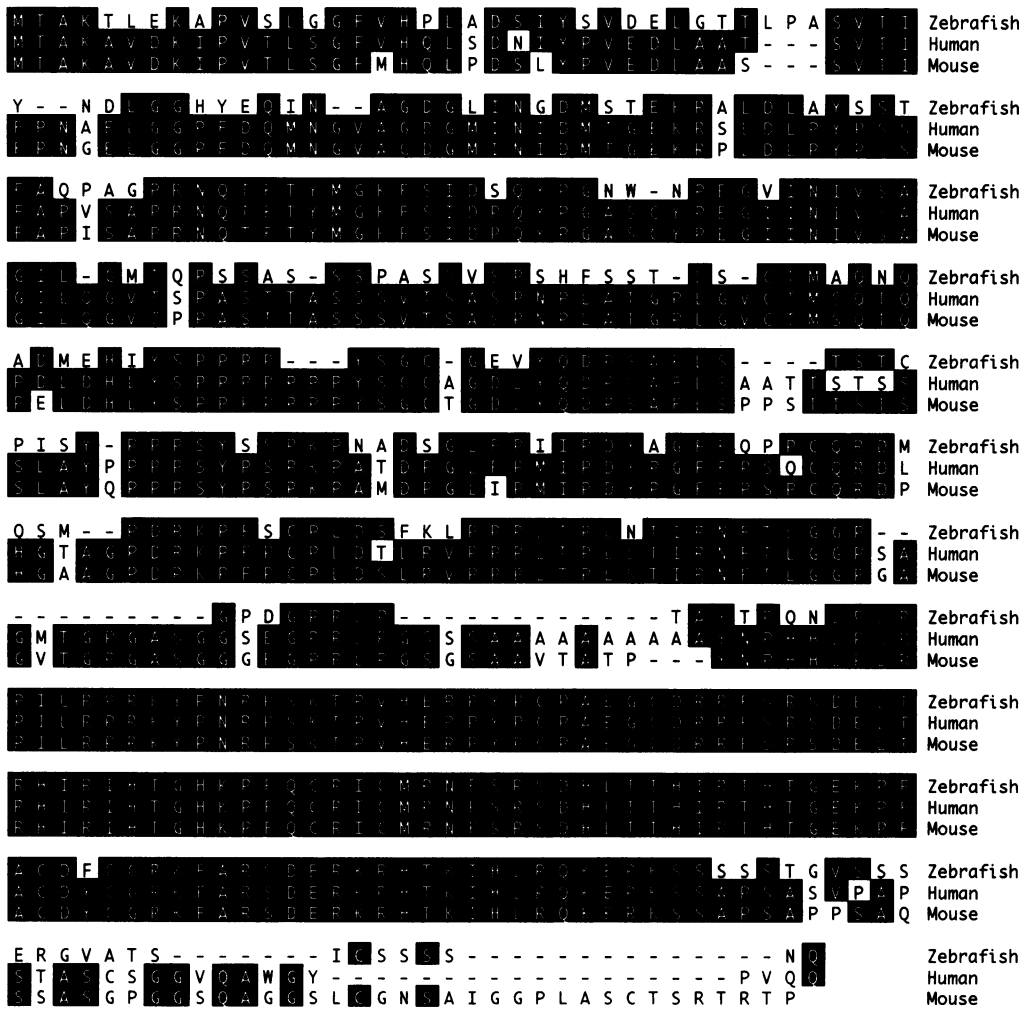


Figure 3. Amino acid sequence comparison of krx-20 with the murine krox-20 (30) and the human (hkrx-20) EGR2 (31). Gaps introduced are represented by dashes. Shaded areas represent regions of sequence identity in all three proteins.

noncoding leader region (35 nucleotides), coding region (1237 nucleotides), and 3′ noncoding trailer region (631 nucleotides). An in-phase TAG stop codon is found 12 nucleotides upstream from the initiator ATG. This first methionine does appear to meet the requirements for the initiation codon (consensus sequence A/GNNATGA/G) as defined by Kozak (26). A potential polyadenylation signal, AATAAA is identified near to the 3′ terminus (underlined in fig. 2A). Upstream from this signal there are two ATTTA sequences (found at positions 1492 and 1506 nucleotides). Multiple corresponding sequences of this type in the 3′ untranslated region of a mRNA have been shown to be related to message instability (27). The presence of a single intron within the coding region was located using synthetic oligonucleotides based on the cDNA sequence (table 1) and a PCR-based strategy. The complete nucleotide sequence of this intron is shown in fig. 2B. The sequence at the 5′ splice site (GG/AGGTAAGG) and that at the 3′ splice site (TTTTCCCTCTTCC/AG) are in good agreement with the splice consensus (28). It is possible that additional introns might be present within the two coding regions of this gene outside the scope of the oligonucleotides. That is within the first 20 amino acids and/or within the last 80 amino acids.

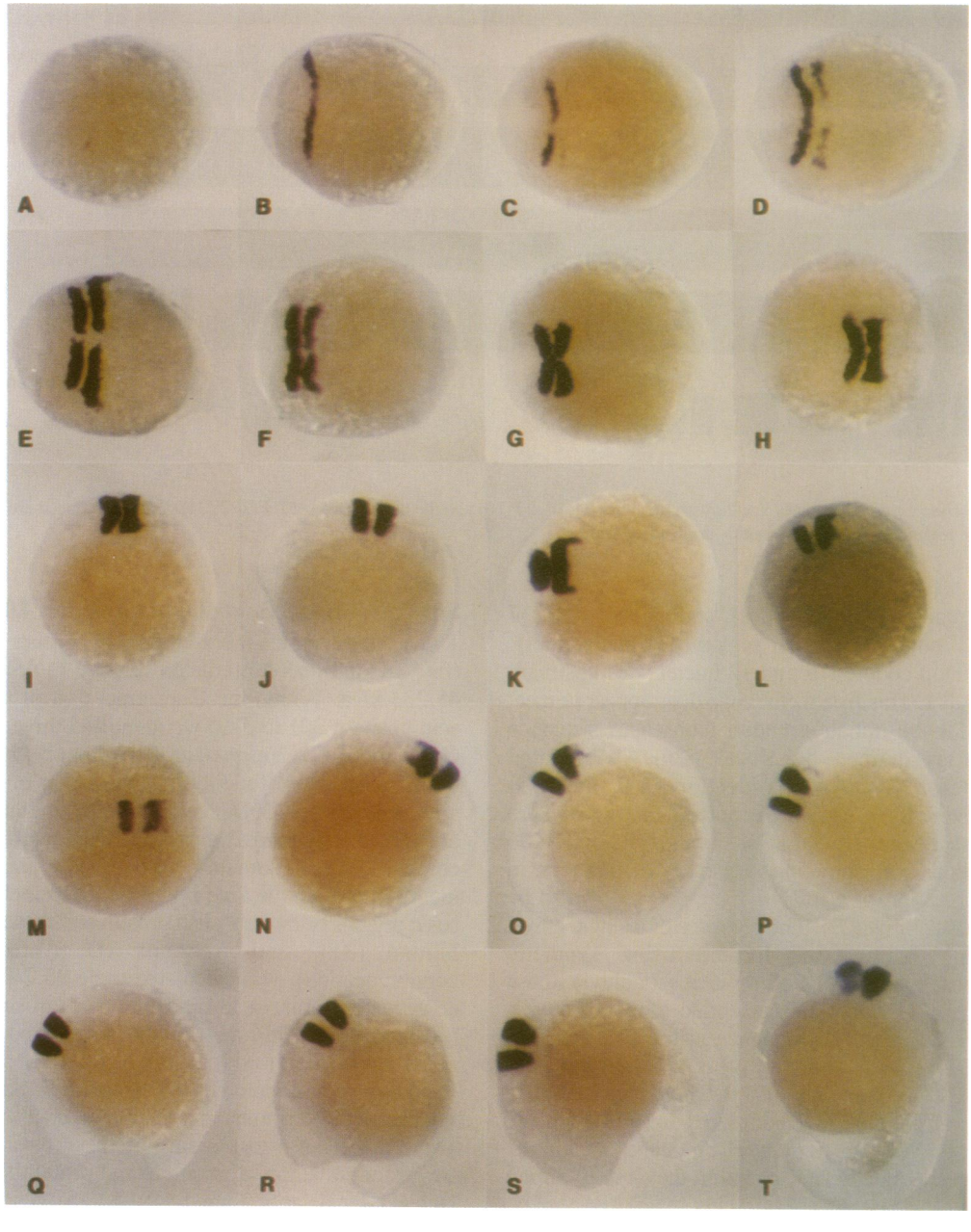
Structural features of the putative krx-20 protein

The deduced zebrafish krx-20 protein consists of 412 amino acids (fig. 2A) and has a predicted relative molecular mass (Mr) of 45.233 kDa. Within the C-terminal part of this protein (amino acid position 301 to 379) we identified three zinc fingers of the C_2-H_2 motif (7). Vesque and Charnay (29) have demonstrated, by using structure function analysis, that only the zinc fingers are necessary for DNA recognition. The zinc fingers of krx-20 are immediately flanked, upstream and downstream by two highly conserved basic domains (fig. 3). The amino acid sequence identity between the zinc finger domains including the flanking basic regions of the zebrafish protein and those of the mouse (30), human (31), and *Xenopus* (D. Wilkinson, personal communication), was calculated to be 98.7%. There is a single conservative amino acid substitution of tyrosine to phenylalanine (position 361). This substitution is observed in both zf187 and zf201 clones therefore it is unlikely that it represents an artifact. In addition conservation of the upstream flanking region has been reported between murine *krox-20*, *EGR-3* and *krox-24* (30, 9, 14). Further, a possible role of these basic regions has been suggested to be the targeting of the protein to the nucleus (29). Nuclear localisation signals (NLS) are generally short (8–10

amino acids) and contain a high proportion of positively charged amino acids. No consensus has yet been derived (32). Within the N-terminal half of krx-20 two acidic regions were identified (positions 22 to 61 and 153 to 173). This result is consistent with the two acidic regions reported for the murine *krox-20* (29). Amino acid comparisons between the three proteins (fig. 3) show that although there is about 50% divergence outside the zinc finger area all the negative charges are conserved. In *Xenopus* the negative charge at position 49 is not conserved (data not shown). This conservation supports the suggestion that these acid regions are transcriptional activation domains.

Finally, the krx-20 protein is very rich in proline residues (12%), though not in the zinc finger domain (4%), and serine residues (13%). The characteristic seven proline residue stretch observed in the mouse is reduced to four residues in the zebrafish (position 160 to 163). Nonetheless three regions were identified where the proline content was around or above 30% and two regions where the serine content was above 50% (positions 184 to 197, 226 to 246, 258 to 302; 124 to 145, 387 to 410 respectively).

The amino acid sequence homology of the entire krx-20 protein to mouse, human and *Xenopus krox-20* was calculated to be 72.0%, 76.4% and 62.0% respectively. Because of the high degree of sequence identity and the segmented pattern of expression during hindbrain development (described below) we conclude that we have isolated the zebrafish homologue of the murine *krox-20* and not that of the closely related *krox*-24 gene (9). *krox-24* has six amino acid differences in the zinc fingers (positions: 294 = proline; 300 = alanine; 303 = valine; 305 = serine; 326 = glutamine; and 361 = isoleucine). These positions refer to the krx-20 protein fig. 2A. In the zebrafish protein one of these six amino acids, that at position 361, differs



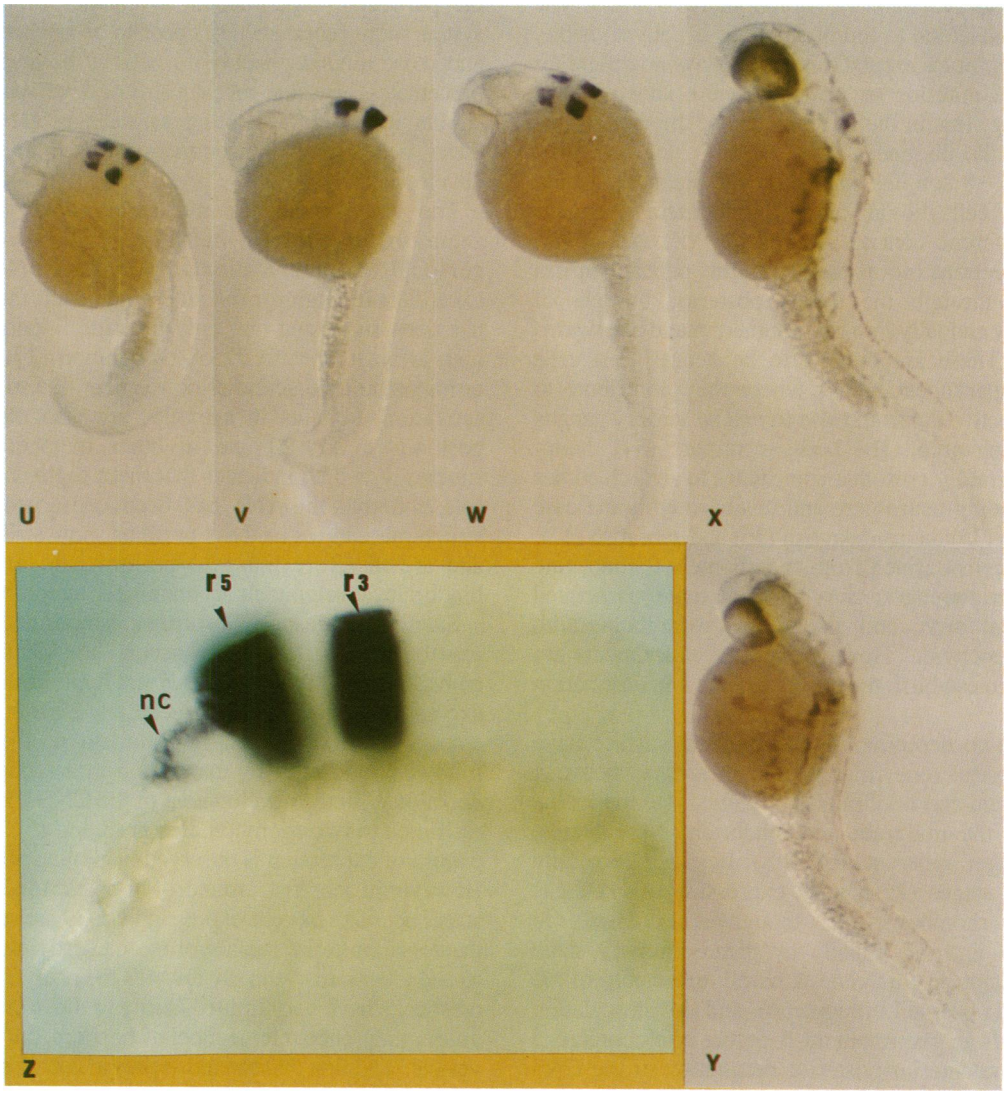


Figure 4. *In situ* hybridisation of *krx-20* on whole mount zebrafish embryos at representative stages of development. **A**–**D**: 100% epiboly, dorsal view. The embryos are oriented with the anterior end to the left. **E**–**H**: Segmentation stage, dorsal view. Morphogenetic movements along the embryonic axis continue. The two domains of the neuroepithelium expressing *krx-20* constrict as segmentation progresses along the neuraxis. **I** and **J**: Neurogenesis, dorsal and side view of the embryos. Anterior end is shown to the left. The transcripts are now restricted in two rhombomeres(r), r3 and r5. **K**–**Q**: Dorsal and side views of embryos between 5–8 somites stage showing neural crest migration from r5. **R**: 14 somites stage. Neural crest migration is no longer apparent. **X**: Primordium 5 stage. Expression in r3 is downregulated but still persists in r5. **Y**: Primordium 11. Expression is barely detectable in the hindbrain. **Z**: Differential interference contrast microscopy showing the same embryo as in **P**, except that here the embryo is viewed from the opposite side (anterior end is to the right). Neural crest (nc) cells appear to be migrating from the dorsal posterior surface of r5 firstly caudally and then ventrally.

from both *krox-20* and *krox-24* (already mentioned above). The remaining five are identical to those of *krox-20*.

Expression of *krx-20* from late gastrulation through straightening

The developmental profile of the expression of *krx-20,* illustrated in figure 4, was produced using *in situ* hybridisation on whole mount zebrafish embryos. Transcripts of this gene are first detectable during late gastrulation, around 100% epiboly, in the prospective neuroepithelium domain. An anterior stripe appears first, followed very soon after by a second (posterior) stripe. These appear to be slightly 'V' shaped pointing caudally and the two arms do not meet at the midline thus leaving a gap where no expression is seen (A–D). Therefore segmentation in the zebrafish hindbrain at the molecular level, occurs around 100% epiboly. This is before the morphological manifestation of neuromeres. The first somite used as the landmark for the beginning of the segmentation stage is not yet visible. This observation is consistent with the suggestion that this gene plays a direct role in the process of segmentation (7).

By the end of gastrulation (around 10 hr postfertilisation) large scale cell movements are nearly completed, except for the developing tail tissue and special cases like the neural crest (33). Morphogenetic movements continue along the embryonic axis and during these, the two *krx-20* transcripts begin to constrict along the neuroaxis (E–H). By 12 hr the presumptive CNS has a tubular appearance (though the division between the brain and spinal cord is not yet evident) and its polarity is marked by the appearance of the optic cap at the rostral end. Zebrafish *krx-20* expression is restricted in two regions of the neuroepithelium of

the prospective hindbrain (I and J), rhombomeres 3 and 5. The otic placode, that lies lateral to rhombomere 5, is not yet visible. A side-view of the embryos around this stage shows no difference in staining, but an anterior or posterior view shows heavy dorsolateral staining, less in the internal part of the neural tube and no staining in the midline.

Around 14 hr (5−8 somites stage, the ear placode still not visible) two tails, of cells moving out of rhombomere 5, appear. Differential interteferene contrast micrоscopy (Z) and *in situ* hybridisations on sections (not presented in this paper) confirm cells moving out through the dorsal posterior borders of rhombomeres firstly caudally into rhombomere 6 and afterwards ventrally (K−Q). These are thought to be neural crest cells migrating to the pharyngeal arches where they contribute to cartilages and connective tissue and also to cranial sensory ganglia associated with each arch. The body metamere (34) theory predicts that premigratory crest that contributes to an arch comes from the corresponding neuromere and this is roughly the case in the avian embryo (35) the medaka fish (36) and zebrafish (37). For example crest cells derived from rhombomere 2 contribute specifically, to the trigeminal sensory ganglion that is associated with the mandibular arch, and also to the mandibular arch cartilages and mesenchyme. However further experiments are necessary in order to establish the exact purpose and destination of these cells.

By about 16 hr two depressions on the dorsal surface have appeared. The rostral one marks the boundary between presumptive forebrain and midbrain. The caudal one marks the boundary between the midbrain and hindbrain. Neural crest migration is no longer apparent (R and S). Between 26 somite and primordium 5 stages (26 hr) the expression of *krx-20* is downregulated in rhombomere 3 though still persists in rhombomere 5 (W and X). Finally by primordium-11 stage (28−30 hr) expression of this gene is barely detectable in the hindbrain (Y). The observed upregulation and downregulation of *krx-20* transcripts in rhombomere 3 before rhombomere 5 probably reflects the anterior-posterior maturation of the CNS and confirms earlier similar observations in the mouse.

CONCLUSION

We have isolated two overlapping cDNAs from zebrafish and these were shown to represent the fish homologue of the murine *krox-20*. The zebrafish gene is designated *krx-20*. In our view, the cloning of developmental genes in zebrafish, that are already known in mammals and *Drosophila*, presents three areas of interest. (1) Firstly, comparing these genes and determining the degree of conservation of their protein structures, functional organization and regulation, throughout evolution. (2) Secondly, establishing and comparing the patterns of expression of related developmental genes in vertebrates. (3) Thirdly, use the zebrafish as a model system to dissect the function of these genes and to study the proteins they produce.

The striking similarities between zebrafish development and that of other vertebrates, together with the accessibility of the zebrafish embryo for observations and manipulations make it an attractive species for studying the mechanisms of development. Further, analysis is facilitated by established methods for generating and characterising mutants (20,21) and production of stable transgenic lines (22,23). Although mutants exhibiting both cell specific defects (37) and disruption of mesodermal segmentation (38) have been recovered, the number is still very small as indeed is the number of cloned genes. Thus despite of the particular experimental advantages of a zebrafish model system more zebrafish genes need to be isolated and characterized and more mutant phenotypes need to be accurately described. Nonetheless, because mechanisms of development are extremely similar in vertebrates (39) resemblances are likely to occur if these genes and their functions are indeed important in development.

The conservation of the amino acid sequence between zebrafish, mouse (30), human (31), *Xenopus* (D. Wilkinson, personal communication) and chick (40), krox-20, is exceptionally high for the zinc fingers and flanking domains. It has been suggested that only the zinc fingers are required for high affinity, specific DNA recognition (29). Towards the N-terminus the two acidic regions, identified to be transcriptional activation domains in krox-20, are also observed at similar positions in krx-20 and in other members of the *krox-20* subfamily, which indicates that these regions have an important role and that this role has been conserved. In addition the structural features of these proteins show striking resemblances and are therefore likely to perform similar functions during hindbrain development in mammals, amphibians and fish.

Such a suggestion is further supported by the fact that expression of this gene occurs at identical sites during embryogenesis. Wholemount *in situ* hybridization established that expression of the zebrafish gene *krx-20* first appears at 100% epiboly as a single anterior domain of the neuroepithelium followed soon after by expression in a second more posterior domain. With the development of the CNS *krx-20* transcripts are restricted to two rhombomeres(r) r3 and r5. This characteristic pattern of expression is observed in all the species so far studied. However in zebrafish around 14 hr postfertilisation (5−8 somites stage) neural crest cells expressing *krx-20* are seen migrating from the dorsal posterior surface of r5 and firstly moving caudally into r6 and afterwards ventrally towards the pharyngeal arches. Neural crest cells have very limited ability to move rostrally or caudally unless they encounter epithelial barriers, in this case, the otic vesicle. The first and second arches in zebrafish are incorporated into jaw structures while the more posterior arches make the definitive gills (46). Rhombomere 3 did not appear to produce migrating neural crest cells, though we do not know whether crest cells are produced and rapidly die or whether no crest cells are formed in this region at all.

Embryonic manipulations, essentially in the avian embryo, have demonstrated the paramount importance of the hindbrain and the neural crest derived from it, in patterning the head region (41). The neural crest seems to be specified before migration as to the structures it will form. It also controls structures formed by other non-neural crest tissues (42). It has been suggested that neural crest cells are patterned according to their rhombomeric origin and during migration the AP order in which they arise is maintained therefore the segmented unit is extended to regions outside the neuroectoderm (43). Lineage tracing experiments in chick embryos were used to define the fate of crest cells from different AP regions of the hindbrain (44). The crest of r2, r4 and r6 mainly populate 1, 2, and 3 branchial arches. This suggests that the developmental processes of the hindbrain and branchial arches are linked (45). Further this association between branchial arches and specific rhombomeres is conserved in all vertebrates at similar stages of development (44,46,47). Extensive mixing between different migrating neural crest populations does not seem to take place which lead to the suggestion that the relative spatial positions and hence the pattern of gene expression of these cells are maintained in the arch they migrate. This observation

is supported by the analysis of the expression patterns of Hox2 genes (48). It was shown that neural crest arising from the hindbrain carries a specific Hox2 positional code (specified according to its rhombomere of origin), which it maintains during its migration to the corresponding branchial arch.

The above studies in chick and mouse suggest that neither r3 nor r5 produce migrating neural crest (44,48). The work presented in this paper, however, demonstrates that this is not the case in zebrafish embryos, at least with reference to r5. In zebrafish crest from r5 will most probably populate the third pharyngeal arch contributing to the cartilages and connective tissue of the definitive gills and possibly to the sensory ganglia associated with this arch. A series of zebrafish coronal sections hybridized with *krx-20* would be essential in order to determine and follow the expression of this gene and the fate of the migrating neural crest cells outside the hindbrain region.

Circumstantial evidence such as different cellular properties in odd and even rhombomeres (49), the alternating pattern of expression of *krox-20* appearing before morphological segmentation, as well as its expression partially overlapping with that of several homeobox-containing genes (50) lead to the suggestions that *krox-20,* might play a role in the regulation of hindbrain segmentation, might act in combination with the hox network to specify odd and even rhombomeres, and might participate in the control of the expression of some of the homeobox containing genes (51). Further, this role is likely to have been conserved throughout evolution since zebrafish, frog, human, mouse and chick homologues of the gene have been cloned and shown to have a segmented pattern of expression. It must be however emphasised that it could be possible for homologous genes to have different functions in different animals as illustrated by the *engrailed* gene (*en*). In *Drosophila* , *en* is part of the gene network that establishes segmentation itself while later it is required for the development of the nervous system (52,53). In zebrafish, *engrailed*, *eng*, expression begins only after the segments appear and therefore cannot be having an equivalent role regarding embryonic segmentation (54). Direct evidence regarding the function of the zebrafish *gene* krx-20 protein must therefore await the outcome of mutagenesis experiments.

ACKNOWLEDGEMENTS

We thank Dr D.Wilkinson for providing us with the murine *krox-20* probe and Dr P.Ingham for support and encouragement during this work; M.Beck, S.Foster and E.Trevarrow for technical help. This work was supported by the Wellcome Trust.

REFERENCES

1. Eisen, J.S. (1991) *The Journal of Neuroscience* **11**, 311–317.
2. Lumsden, A. and Keynes, R. (1989) *Nature* **337**, 424–428.
3. Fraser, S., Keynes, R. and Lumsden, A. (1990) *Nature* **344**, 431–435.
4. Garcia-Bellido, A., Ripoll, P. and Morata, G. (1973) *Nature New Biology* **245**, 251–253.
5. Wilkinson, D.G., Bhatt, S., Chavrier, P., Bravo, R. and Charnay, P. (1989) *Nature* **337**, 461–464.
6. Wilkinson, D.G. (1990) *Developmental Biology* **1**, 127–134.
7. Chavrier, P., Zerial, M., Lemaire, P., Almendral, J., Bravo, R. and Charnay, P. (1988) *EMBO J.* 7, 29–35.
8. Chavrier, P., Vesque, C., Galliot, B., Vigneron, M., Dollé, P., Duboule, D. and Charnay, P. (1990) *EMBO J.* **9**, 1209–1218.
9. Lemaire, P., Revelant, O., Bravo, R. and Charnay, P. (1988) *Proc. Nat. Acad. Sci. USA* **85**, 4691–4695.
10. Sukhatme, V.P., Cao, X., Chang, L.C., Tsai-Morris, C.H., Stamenkovich, D., Ferreira, P.C.P., Cohen, D.R., Edwards, S.A., Shows, T.B., Curran, T., Le Beau, M.M. and Adamson, E.D. (1988) *Cell* **53**, 37–43.
11. Christy, B.A., Lau, L.F. and Nathans, D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7857–7861.
12. Milbrandt, J. (1987) *Science* **238**, 797–799.
13. Lim, R.W., Vanum, B.C. and Herschman, H.R. (1987) *Oncogene* **1**, 263–270.
14. Patwardhan, S., Gashler, A., Siegel, M.G., Chang, L.C., Joseph, L.J., Shows, T.B., Le Beau, M.M. and Sukhatme, V.P. (1991) *Oncogene* **6**, 917–928.
15. Crosby, S.D., Puetz, J.J., Simburger, K.S., Fahrner, T.J. and Milbrandt, J. (1991) *Mol. Cell. Biol.* **11**, 3835–3841.
16. Wieschaus, E., Nüsslein-Volhard, C. and Kluding, H. (1984) *Devel. Biol.* **104**, 172–186.
17. Tautz, D., Lehmann, R., Schnürch, H., Schuh, R., Seifert, E., kielin, A., Jones, K. and Jäckle, H. (1987) *Nature* **327**, 383–389.
18. Vincent, A., Colot, H.V. and Rosbash, M. (1985) *J. Mol. Biol.* **186**, 149–166.
19. Mével-Ninio, M., Terracol, R. and Kafatos, F.C. (1991) *EMBO J.* 10, 2259–2266.
20. Streisinger, G., Walker, C., Dower, N., Knauber, D. and Singer, F. (1981) *Nature* **291**, 293–296.
21. Chakrabarti, S., Streisinger, G., Singer, F. and Walker, C. (1983) *Genetics* **103**, 109–123.
22. Stuart, G.W., Vielkind, J.R., McMurray, J.V. and Westerfield, M. (1990) *Development* **109**, 577–584.
23. Culp, P., Nüsslein-Volhard, C. and Hopkins, N. (1991) *Proc. Natl. Acad. Sci. USA* 88, 7953–7957.
24. Lorens, J.B. (1991) *PCR Methods and Applications* 1, 140–141.
25. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *In* Molecular Cloning, A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).
26. Kozak, M. (1987) *Nucl. Acids Res.* **15**, 8125–8148.
27. Shaw, G. and Kamen, R. (1986) *Cell* **46**, 659–667.
28. Shapiro, M.B. and Senapathy, P. (1987) *Nucl. Acids Res.* **15**, 7155–7174.
29. Vesque, C. and Charnay, P. (1992) *Nucl. Acids Res.* **20**, 2485–2492.
30. Chavrier, P., Janssen-Timmen, U., Mattéi, M.G., Zerial, M., Bravo, R. and Charnay, P. (1989) *Mol. Cell. Biol.* **9**, 787–797.
31. Joseph, L.J., Le Beau, M.M., Jamieson, Jr., G.A., Archarya, S., Shows, T.B., Rowley, J.D. and Sukhatme, V.P. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7164–7168.
32. Garcia-Bustos, J., Heitman, J. and Hall, M.N. (1991) *Biochim. Biophys. Acta* 1071, 83–101.
33. Kimmel, C.B., Warga, R.M. and Schilling, T.F, (1990) *Development* **108**, 581–594.
34. Goodrich, E.S. (1930) Studies on the Structure and Development of Vertebrates, MacMillan, London.
35. Couly, G. and Le Douarin, N.M. (1990) *Development* **108**, 543–558.
36. Langille, R.M. and Hall, B.K. (1988) *Anat. Embryol.* 177, 297–305.
37. Grunwald, D.J., Kimmel, C.B., Westerfield, M., Walker, C. and Streisinger, G. (1988) *Devel. Biol.* **126**, 115–128.
38. Kimmel, C.B., Kane, D.A., Walker, C., Warga, R.M. and Rothman, M.B. (1989) *Nature* **337**, 358–362.
39. Kimmel, C.B. (1989) *Trends in Genetics* 5, 283–288.
40. Nieto, M.A., Bradley, L.C. and Wilkinson, D.G. (1991) *Development Supplement* **2**, 59–62.
41. Le Douarin, N.M. (1982) *The Neural Crest.* Cambridge Univ. Press, Cambridge, UK.
42. Noden, D.M. (1986) *Developmental Biology* **116**, 347–356.
43. Couly, G. and Le Douarin, N.M. (1990) *Development* 108, 543–558.
44. Lumsden, A., Sprawson, N. and Graham, A. (1991) *Development* **13**, 1281–1291.
45. Lumsden, A. (1990) In *Seminars in Developmental Biology: The Evolution of Segmental Patterns* **1**, 117–125.
46. Kimmel, C., Schilling, T. and Hatta, K. (1991) *Current Topics in Develop. Biol.* **25**, 77–110.
47. Kuratani, S. and Kirby, M. (1991) *Am. J. Anat.* **191**, 215–227.
48. Hunt, P., Wilkinson, D. and Krumlauf, R. (1991) *Development* 112, 43–50.
49. Lumsden, A. and Guthrie, S. (1991) *Development Supplement* **2**, 9–15.
50. Wilkinson, D.G., Bhatt, S., Cook, M., Boncinelli, E. and Krumlauf, R. (1989) *Nature* 341, 405–409.
51. Gilardi, P., Schneider-Maunoury, S. and Charnay, P. (1991) *Biochimie* 73, 85–91.
52. Lawrence, P.A. and Struhl, G. (1982) *EMBO J.* **1**, 827–833.
53. Lawrence, P.A. and Johnston, P. (1984) *EMBO J.* **3**, 2839–2844.
54. Patel, N.H., Martin-Blanco, E., Coleman, K.G., Poole, S.J., Ellis, M.C., Kornberg, T.B. and Goodman, C.S. (1989) *Cell* **58**, 955–968.